

ANALISIS KEPUASAN PENGGUNAAN APLIKASI SHOPEE MENGUNAKAN ALGORITMA NAÏVE BAYES

Gabriel Carvenita Triasis¹⁾ Desi Arisandi²⁾ Tri Sutrisno³⁾

¹⁾²⁾³⁾ Teknik Informatika, FTI, Universitas Tarumanagara
Jl. Letjen S Parman No 1, Jakarta 11440 Indonesia
email : gabriel.535180124@stu.untar.ac.id¹⁾, email : desia@fti.untar.ac.id²⁾

ABSTRACT

The Covid-19 pandemic that attacks the world's health conditions, including Indonesia, greatly affects daily life. The impacts include the need to maintain distance and avoid crowds, which forces people to reduce interactions between people. Therefore, people need other ways to be able to meet their daily needs. One alternative that can be done is to use e-commerce applications to conduct online transactions or also known as online shopping. It was reported on September 30, 2021 that shopee occupies the most superior e-commerce in Indonesia with the number of monthly visitors which currently reaches 93 million subscribers, followed by Tokopedia with 86 million subscribers and in third place with 35 million customers occupied by Bukalapak. Therefore, an analysis of satisfaction with the use of e-commerce applications with the Shopee case study will be carried out to determine the response from users to the applications that are declared the most desirable. The application will be made based on a website to make it easier to use and use the Naïve Bayes algorithm as the analysis method. The Naïve Bayes algorithm is one method that can classify data which in this case are comments from application users into three types of comments, namely positive, negative and neutral. The process that is run also includes a preprocessing task as raw data processing into ready-to-use data and a confusion matrix as a calculation of the accuracy of the resulting application. Testing of the application is carried out in two ways, namely using a confusion matrix using a system with 80% results and a user acceptance test with 94.8% results, namely very good predicate. From the analysis, it can be concluded that Shopee users are satisfied with the application used and the classification of user reviews using the Naïve Bayes Algorithm produces a fairly good accuracy.

Key words

Algoritma Naïve Bayes, Analisis Kepuasan Penggunaan , Analisis Sentimen

1. Pendahuluan

Pandemi Covid-19 yang menyerang kondisi kesehatan dunia termasuk Indonesia sangat mempengaruhi kehidupan sehari-hari. Dampak yang ada diantara lain adalah perlunya menjaga jarak dan menghindari kerumunan yang mana memaksa masyarakat untuk mengurangi interaksi antar sesama. Oleh karena itu masyarakat memerlukan cara lain untuk dapat memenuhi kebutuhan sehari-hari. Salah satu alternatif yang dapat dilakukan adalah dengan menggunakan aplikasi *e-commerce* untuk melakukan transaksi *online* atau juga dikenal sebagai *online shopping*.

Dilansir pada tanggal 30 September 2021 bahwa shopee menduduki *e-commerce* paling unggul di Indonesia dengan jumlah pengunjung bulanannya yang saat ini mencapai 93 juta pelanggan diikuti dengan peringkat kedua yaitu Tokopedia dengan 86 juta pelanggan dan di peringkat ketiga dengan 35 juta pelanggan diduduki oleh Bukalapak. Maka dari itu analisis kepuasan penggunaan aplikasi *e-commerce* dengan studi kasus *Shopee* akan dilakukan untuk mengetahui respon dari pengguna terhadap aplikasi yang dinyatakan paling diminati tersebut.

Adapun aplikasi akan dibuat berbasis *website* untuk memudahkan penggunaan dan menggunakan algoritma Naïve Bayes sebagai metode analisisnya. Algoritma Naïve Bayes merupakan salah satu metode yang dapat mengklasifikasi data yang dalam kasus ini merupakan komentar dari pengguna aplikasi menjadi tiga jenis komentar yaitu positif, negatif dan netral. Proses yang dijalankan juga mencakup *preprocessing task* sebagai pengolahan data mentah menjadi data siap pakai dan *confusion matrix* sebagai perhitungan akurasi dari aplikasi yang dihasilkan. Pengujian terhadap aplikasi dilakukan dengan dua cara yaitu menggunakan *confusion matrix* menggunakan sistem dengan hasil 80% dan *user acceptance test* dengan hasil 94,8% yaitu predikat *Sangat Baik*. Dari analisis yang dilakukan dapat

disimpulkan bahwa pengguna *Shopee* puas dengan aplikasi yang digunakan dan pengklasifikasian *review* pengguna menggunakan Algoritma Naïve Bayes menghasilkan akurasi yang cukup baik.

2. Dasar Teori

Dalam melakukan perancangan sistem, diperlukan landasan teori sebagai acuan dan panduan terkait rancangan yang dibuat. Berikut teori yang digunakan antara lain Analisis Sentimen, Algoritma Naïve Bayes, *Preprocessing Task* dan *Confusion Matrix*.

2.1 Analisis Sentimen

Analisis sentimen atau *sentiment analysis* merupakan proses penggunaan *text analytics* untuk mendapatkan data dari internet dan media sosial. Analisis sentimen bertujuan untuk memperoleh opini atau pendapat dari pengguna yang terdapat pada *platform* terkait. [1] Komentar atau ulasan dari sebuah aplikasi merupakan salah satu sumber pengekspresian atau implementasi dari analisis sentimen. Informasi yang disajikan biasanya berbentuk teks atau gambar dalam format forum, blog, media sosial atau aplikasi dan situs berisi *review* dari pengguna. Dengan analisis sentimen, informasi tersebut yang semula tidak terstruktur dapat diubah menjadi data yang lebih terstruktur.

2.2 Algoritma Naïve Bayes

Algoritma Naïve Bayes adalah salah satu algoritma klasifikasi berdasarkan teorema Bayesian pada statistika (Suntoro, Wahyu, & Indriyawati, 2018). Algoritma Naïve Bayes dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas (Han and Kamber, 2012). [2]

Teorema Bayesian menghitung nilai *posterior probability* $P(H|X)$ menggunakan probabilitaas $P(H)$, $P(X)$, dan $P(X|H)$ (Kantardzic, 2011), dimana nilai X adalah data testing yang kelasnya belum diketahui. Nilai H adalah hipotesis data X yang merupakan suatu kelas yang lebih spesifik. Nilai $P(X|H)$ atau disebut dengan *likelihood* adalah probabilitas hipotesis X berdasarkan kondisi H . Nilai $P(H)$ atau disebut juga dengan *prior probability* adalah probabilitas hipotesis H . sedangkan nilai $P(X)$ yang disebut juga dengan *predictor prior probability* adalah probabilitas X .

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Keterangan:

X : data dengan class yang belum diketahui

H : hipotesis data berupa suatu class spesifik

$P(H|X)$: probabilitas hipotesis H berdasar kondisi X

$P(H)$: probabilitas hipotesis H

$P(X|H)$: probabilitas X berdasarkan kondisi pada hipotesis H
 $P(X)$: probabilitas X

Setelah itu hitung himpunan kelas dengan rumus sebagai berikut:

$$V_{map} = \underset{v_j \in V}{argmax} \frac{P(a_1, \dots, a_n | v_j) P(v_j)}{P(a_1, \dots, a_n)} \quad (2)$$

Setelah diperoleh perhitungan, maka kelas yang didapat adalah kelas dengan nilai tertinggi. Nilai $P(V_j)$ ditentukan berdasarkan persamaan berikut:

$$P(V_j) = \frac{|banyak\ kata\ dalam\ kelas|}{|banyak\ kata|} \quad (3)$$

$$P(W_k | V_j) = \frac{n_k + 1}{n + |kata|} \quad (4)$$

Keterangan:

$P(W_k | V_j)$: probabilitas kata W_k muncul dengan kelas V_j

n_k : frekuensi kata W_k muncul dengan kelas V_j

n : banyak kata pada kelas V_j

$|kata|$: banyak kata pada data

Langkah-langkah algoritma Naïve Bayes adalah sebagai berikut:

1. Siapkan dataset
2. Hitung jumlah kelas pada data training
3. Hitung jumlah kasus yang sama dengan kelas yang sama.
4. Kalikan semua hasil sesuai dengan data testing yang akan dicari kelasnya.
5. Bandingkan hasil perkelas, nilai tertinggi ditetapkan sebagai kelas baru.

2.3 Preprocessing Task

Preprocessing task atau *text preprocessing* merupakan suatu proses untuk menyeleksi data text yang semula tidak terstruktur menjadi lebih terstruktur dengan melalui serangkaian tahapan.[3] Tahapan tersebut meliputi *case folding*, *tokenizing*, *filtering*, *stemming* dan normalisasi. Berikut penjelasan terkait tahapan yang ada:

1. Case Folding

Peran *case folding* pada pemrosesan teks ini adalah mengkonversi keseluruhan teks dalam kalimat tersebut untuk menjadi kalimat dengan bentuk yang standar.

2. Tokenizing

Tahapan pemotongan *string input* berdasarkan tiap kata yang terdapat didalamnya.

3. Filtering

Tahapan ini merupakan tahap mengambil kata penting yaitu penghilangan *stopwords* seperti "yang", "dan", "di", "dari" dan yang lainnya.

4. Stemming

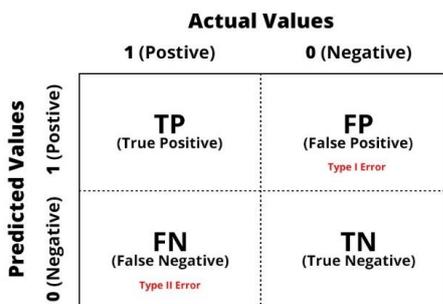
Stemming merupakan proses menghilangkan imbuhan seperti ter-, me-, -nya, di- dan lain sebagainya.

5. Normalisasi

Normalisasi adalah proses penghilangan kata berlebih dan menormalisasikan data menjadi bentuk utuh dengan tidak mengandung singkatan sehingga kata yang dihasilkan merupakan kata inti dari kalimat dengan struktur yang lebih rapi.

2.4 Confusion Matrix

Confusion matrix adalah pengukuran performa untuk masalah klasifikasi machine learning dimana hasil yang didapat berupa dua kelas atau lebih. [4] Pada dasarnya, confusion matrix memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi sebenarnya. Terdapat empat istilah sebagai representasi hasil proses klasifikasi pada confusion matrix diantaranya adalah True Positif (TP), True Negatif (TN), False Positif (FP) dan False Negatif (FN) seperti yang dapat dilihat pada gambar berikut.

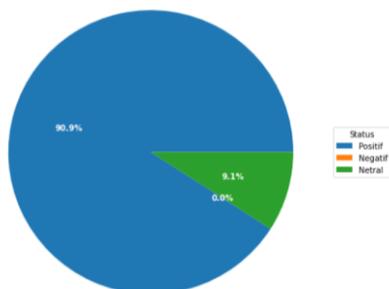


Gambar 1 Tabel Confusion Matrix

3. Hasil Percobaan

3.1 Pengujian Terhadap Metode

Hasil yang didapat melalui aplikasi merupakan akurasi sebesar 90% dengan grafik seperti pada gambar berikut.



Gambar 2 Hasil Perhitungan Confusion Matrix Menggunakan Sistem

Sedangkan untuk perhitungan manual menyatakan bahwa tingkat akurasi sebesar 80% dengan perhitungan sebagai berikut:

$$Accuracy = \frac{8}{10} \times 100\% = 80\% \quad (5)$$

Hal ini dikarenakan perbedaan data klasifikasi antara pengklasifikasian otomatis dengan manual yang mana pada data klasifikasi otomatis dengan data klasifikasi Naive Bayes hanya terdapat satu perbedaan klasifikasi atau satu false condition sedangkan dengan menggunakan klasifikasi manual terdapat dua false condition yang didapat.

3.2 Pengujian Terhadap Rancangan

Hasil pengujian user acceptance test merupakan pengujian menggunakan google form dengan total responden sebanyak 41 user yang sudah dipastikan berbeda dikarenakan form sudah menggunakan setting sekali pakai. Hasil yang didapat adalah sebagai berikut:

Tabel 1 Hasil Pengujian User Acceptance Test

Pertanyaan	Persentase
Apakah website tersebut mudah dipahami?	SB
Apakah tampilan yang diberikan nyaman untuk digunakan?	SB
Dari data yang dihasilkan, apakah anda dapat memahami hasil dari analisis?	SB
Apakah hasil analisis tersebut merupakan data yang informatif?	SB
Setelah melihat hasil analisis, apakah anda setuju bahwa pengguna aplikasi Shopee merasa puas dengan aplikasi tersebut?	SB
Apakah setelah melihat hasil analisis, anda tertarik untuk menggunakan aplikasi Shopee?	SB

4. Kesimpulan

Setelah proses pengujian yang dilakukan terdapat kesimpulan yang dapat dirangkum menjadi beberapa poin yakni sebagai berikut:

1. Analisis kepuasan penggunaan aplikasi menggunakan data berupa review dari pengguna dapat dinyatakan berhasil dikarenakan setelah melakukan proses analisis baik secara sistem maupun manual dapat ditemukan bahwa review yang diberikan oleh pengguna yang sudah diklasifikasikan menggunakan Algoritma Naïve Bayes memiliki hasil yang

menyatakan bahwa pengguna puas dengan aplikasi tersebut.

2. Penggunaan Algoritma Naïve Bayes sebagai metode klasifikasi memiliki tingkat akurasi yang cukup tinggi namun tingkat akurasi tersebut tidak stabil yang mana tidak dapat dipastikan untuk setiap percobaan dapat menghasilkan persentase yang berbeda-beda. Namun untuk keseluruhan persentase rata-rata dari beberapa percobaan yang dijalankan, persentase yang dihasilkan mencapai angka lebih dari 70%.
3. Pengujian metode perhitungan dengan menggunakan *confusion matrix* sangat dianjurkan untuk dilakukan secara 2 cara yaitu secara sistem dan secara manual. Hal ini untuk mengetahui apakah setelah melalui 2 cara tersebut metode yang digunakan masih dapat mencapai persentase yang baik. Dalam pengujian aplikasi Analisis Kepuasan Penggunaan Aplikasi *E-commerce* dengan studi kasus *Shopee* mencapai tingkat akurasi sebesar 80% dengan perhitungan manual dan 90% dengan perhitungan secara sistem. Dari hasil ini dapat dinyatakan bahwa aplikasi ini memiliki persentase pengujian yang baik.
4. Berdasarkan hasil pengujian *user acceptance test* terhadap 41 responden dapat dinyatakan bahwa aplikasi ini merupakan aplikasi yang baik dan dapat nyaman digunakan oleh *user* serta memenuhi tujuan dari pembuatan aplikasi ini sendiri.

REFERENSI

- [1] Domi Sepri, 2020, “Penerapan Algoritma Naïve Bayes Untuk Analisis Kepuasan Penggunaan Aplikasi Bank”, Padang.
- [2] Fajar Ratnawati, 2018, “Implementasi Algoritma Naïve Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter”, Riau.
- [3] Navi Atri Lestari, Tubagus Mohammad Akhriza, dan Eka Yuniar, 2020, “Metode Naïve Bayes Classifier Dengan Textblob Untuk Analisis Sentimen Terhadap Pelayanan Indihome Dan First Media”, Seminar Nasional Teknologi Informasi Dan Komunikasi STI&K (SeNTIK).
- [4] Devin Abipraya, 2021, “Implementasi *Opinion Mining* Untuk *Provider* Internet Menggunakan Metode *Naïve Bayes Classification*”, Jakarta.

Gabriel Carvenita Triasis, mahasiswa S1, program studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Tarumanagara.